Model-related outcome differences in power system models with sector coupling - quantification and drivers

Hans Christian Gils[a,b,*], Hedda Gardian[a,b], Martin Kittel[c], Wolf-Peter Schill[c], Alexander Murmann[d], Jann Launer[e], Felix Gaumnitz[f], Jonas van Ouwerkerk[g,h,i], Jennifer Mikurda[j], Laura Torralba-Díaz[k,b]

[a] *German Aerospace Center (DLR), Institute of Networked Energy Systems, Curiestr. 4, 70563 Stuttgart, Germany*
[b] *Stuttgart Research Initiative on Integrated Systems Analysis for Energy (STRise), Keplerstraße 7, 70174 Stuttgart, Germany*
[c] *German Institute for Economic Research (DIW Berlin), Mohrenstraße 58, 10117 Berlin, Germany*
[d] *Research Center for Energy Economics (FfE), Am Blütenanger 71, 80995 München, Germany*
[e] *Reiner Lemoine Institute, Rudower Chaussee 12, 12389 Berlin, Germany*
[f] *Institute for High Voltage Equipment and Grids, Digitalization and Energy Economics (IAEW), RWTH Aachen University, Schinkelstraße 6 52056 Aachen, Germany*
[g] *Institute for Power Electronics and Electrical Drives (ISEA), RWTH Aachen University, Jägerstraße 17-19, 52066 Aachen, Germany*
[h] *Institute for Power Generation and Storage Systems (PGS), E.ON ERC, RWTH Aachen University, Mathieustraße 10, 52074 Aachen, Germany*
[i] *Jülich Aachen Research Alliance, JARA-Energy*
[j] *Chair for Management Science and Energy Economics (EWL), University of Duisburg-Essen, Universitätsstr. 11, 45117 Essen, Germany*
[k] *Institute of Energy Economics and Rational Energy Use (IER), University of Stuttgart, Heßbrühlstraße 49a, 70565 Stuttgart, Germany*

Abstract

This paper presents the results of a multi-model comparison aimed at determining outcome deviations resulting from differences in power systems models. We apply eight temporally and spatially resolved models to 16 stylized scenarios of the Central European power system. These scenarios differ in variable renewable energy supply share, technology scope, and optimization scope. We focus on technologies for balancing the variability of power generation, such as dispatchable power plants, energy storage, power transmission, and flexibility related to sector coupling. To separate model-related from data-related outcome deviations, we use harmonized input data in all models. We find that our approach allows to isolate and quantify model-related outcome deviations and robust effects with regard to system operation and investment decisions. Furthermore, we can attribute these deviations to the identified model differences. Our results show that trends in the use of individual flexibility options are ro-

*Corresponding author
Email address: `hans-christian.gils@dlr.de` (Hans Christian Gils)

bust across most models. Moreover, we our analysis reveals that differences in the general modeling approach and the modeling of specific technologies lead to comparatively small deviations, whereas a heterogeneous model scope can cause substantially larger deviations. Due to the large number of models and scenarios, our analysis can provide important information on which investment and operation decisions are robust to the model choice, and which modeling approaches have a particularly high impact on results. Our findings may guide both modelers and decision makers in properly evaluating the results of similarly designed power system models.

Word count: 7785

Keywords: power sector modeling, model comparison, sector coupling, optimization

---

Highlights

- Harmonized application of eight power sector models with sector coupling
- Consideration of 16 stylized scenarios differing in complexity and supply structure
- Deviations in fully harmonized models are mostly below 20% and can be traced back
- Different model scopes cause broader variation in the use of flexibility options
- Use of sector coupling for renewable energy integration is robust across all models

List of abbreviations

BEV battery electric vehicle

CHP combined heat and power

CO_2 carbon dioxide

DC direct current

DH district heating

DR demand response

E2P energy to power

HP heat pump

IAM integrated assessment model

PV photovoltaics

TES thermal energy storage

VRE variable renewable energy

1. Introduction

1.1. Background and motivation

Optimizing system models are among the standard tools used for energy system analysis. Such models are often applied for investigating future energy supply systems. When focusing on power supply, the integration of fluctuating power generation from variable renewable energy (VRE) through flexibility options such as storage, grids, and dispatchable power plants is the focus of many models and their application [1]. This is increasingly complemented by analyses of the flexibility that can be tapped when implementing so-called sector coupling [2]. This essentially refers to the direct and indirect use of electricity in other areas of the energy system in order to reduce greenhouse gas emissions there. Of great importance in this context is the partially flexible use of electricity for charging battery electric vehicles (BEVs) [3], for heat generation in heat pumps (HPs) and electric boilers [4], and the electrolytic generation of hydrogen [5].

However, scenarios based on the application of such models often come to very different conclusions regarding future technology use. Differences in model outcomes can result from any of the steps in the modeling chain [6], e.g. from different assumptions regarding the development of demand, costs and technology parameters (data-related differences), but also from different modeling approaches (model-related differences). The fundamental cause of divergent model outcomes is the necessity to abstract complex systems, which can be realized in different ways. The mathematical formulation of the model plays a role here, as do the scope and detail of the spatial, temporal and technological model dimensions. A comprehensive understanding of the effects of particular modeling approaches is thus a prerequisite for the correct interpretation of the scenarios and model results. Structured model comparisons are a helpful tool to gain this understanding. To quantify model-related deviations in outcomes, data-related differences must be kept as small as possible by using harmonized input data [7].

1.2. State of research

The literature offers a range of publications on comparisons of energy or power system models. However, these are mostly limited to a theoretical comparison of the methodology used, the model scope and the model properties. The most recent of these papers focus, for example, on the consideration of policy instruments in the models [8], the technological focus [9], the ability to address policy-relevant research questions [10], the comparison of model resolutions [11], the ability to analyze renewable energy systems [12], or the optimization of multi-energy systems [13]. An application of the respective compared models does not take place in any of these publications.

In contrast, Sugiyama et al. [14] present a comparative application of spatially and temporally aggregated energy system models to transformation scenarios for Japan. However, an input data harmonization, an analysis of model differences and a comparison of model properties were not performed. Similarly, a set of eight models differing widely in their temporal and spatial detail was applied to a scenario analysis for the North-American energy system in [15]. Their comparison is also not based on fully harmonized input data, and the differences in results are not related to model properties. North America is also the assessment area of another comparison considering 17 models and 13 scenarios [16]. Again, there is no harmonization of the input data and no analysis of model-related outcome deviations. The model comparison of Giarola et al. [17] is also devoted to North America, but examines future energy storage expansion in particular. Since the four models used have numerous differences in scope and input data, there are extensive deviations in the results, which can only be partially attributed to model differences.

A coordinated application of four models to three scenarios of a future German power system including flexible sector coupling and regional resolution was carried out in Gils et al. [18]. In Siala et al. [19], a systematic comparison of the effects of different model types, planning horizons, spatial and temporal resolutions was performed applying five models to different power system scenarios for Germany. Both works rely on harmonized model input data, but do not provide a systematic analysis of model-related outcome deviations. The deployment pathways of VRE technologies in the United States of America was the focus of another comparison of three energy planning models [20]. Despite the use of harmonized input data, large ranges of plant expansion result there, which is attributed to different technology modeling of VRE. To explore the effect of model differences in detail, [7] applied nine models with fully harmonized input data to highly simplified use cases. On this basis, the impact of differences in model formulation can be well understood, but there is no transferability to more realistic scenarios.

In the field of integrated assessment modeling, there is a wide range of publications in which multiple models are applied to transformation scenarios. For example, [21] focuses in particular on VRE integration modeling, [22] on energy technology cost assumptions, [23] on carbon price impacts, and [24] on national contributions for the achievement of the Paris agreement. All these works have in common that different models are applied to the respective scenarios considered, but without harmonization of input data and detailed exploration of differences in results. Methods for the harmonization of input data of integrated assessment models (IAMs) are addressed in some recent works. Krey et al. [25] conduct a review on techno-economic parameters in IAMs, and encounters significant differences. They also identify numerical differences in technology modeling as another possible cause of differences. Data harmonization as well as model application is not performed. Giarola et al. [26] address how such a harmonization could be implemented and which challenges would have to be overcome. They show that with the application of the developed framework for data harmonization, the differences in IAM results can potentially be reduced.

What most of these publications have in common, however, is that differences in results are not systematically explored and attributed to model properties. In addition, complete harmonization of input data and model configurations is done only in very few cases. Similarly, model-related differences in results are not captured for individual technologies. Furthermore, the consideration of flexible sector coupling plays no or only a minor role in earlier comparisons of power system models.

1.3. Contribution of this paper

Complementing the existing literature, our paper provides a systematic assessment of model-related differences in power system models, considering stylized future energy scenarios. In particular, we focus on technologies for balancing fluctuating power generation from VRE, referred to as flexibility options. These include electricity storage, transmission grids, and flexible sector coupling. Given this focus, our model comparison includes a portfolio of eight power system models optimizing one year of system operation in an hourly resolution. These models are applied to a total of 16 use cases that differ substantially in their design. While one part of these cases considers a complete harmonization of the technology scope in the models, this is not the case in the other part. Thus, outcome differences can also be correlated with the choice of technology scope. Regardless of the technologies considered in each use case, all models use a harmonized set of input data. The focus of this study is to answer the following research questions:

1. How large are the model-related differences of optimizing power system models with complete harmonization of model scope and input data, and how are they related with the VRE share?
2. Do robust statements on technology deployment result even with different model scopes?
3. Which differences in modeling approach and technology modeling have a particularly strong impact on the composition and operation of the optimal system?

To answer these questions, model differences are first collected and categorized. On this basis, the deviations between the results are then systematically analyzed and correlated with the model differences. The results of our comparison strengthens the understanding of the effect of differences in temporally and spatially resolved power system models. This is equally helpful for developers and users of models, as well as for decision makers who use the model results.

The paper is divided into four main parts. Section 2 sets out the methodology of the model comparison. Based on this, Section 3 presents the modeling results and their analysis. This is followed by a discussion of the results in Section **??**, as well as the derivation of the main conclusions in Section 4.

2. Materials and Methods

2.1. Set-up and input data of the model comparison

The basic approach in the model comparison is essentially characterized by the use of harmonized input data and stylized, yet systematic, use cases. On the data side, we build on a previous model comparison exercise, which focused on technology-specific modeling differences [7]. The harmonized input data set includes all exogenous plant capacities, techno-economic parameters and time series.

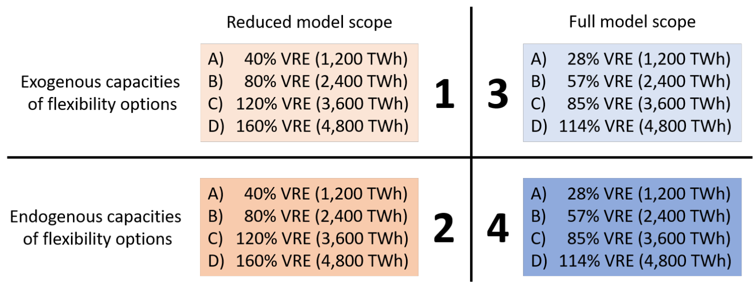

Figure 1: Overview of the use cases considered in the model comparison. Groups 1 to 4 differ in the technology scope and consideration of endogenous capacity optimization of flexibility options. Within each group, four different sets of VRE capacities and thus amounts of renewable electricity are considered. The achievable VRE supply shares differ due to the different demands depending on the model scope.

The 16 use cases considered differ in three characteristics: the VRE capacities, the consideration of an endogenous capacity expansion of flexibility options, and the technology scope (Figure 1). The technology scope correlates with the degree of harmonization of the models. In the case of a reduced consideration (use cases 1 and 2) the same technologies are considered in all models, whereas in the case of a more extensive consideration (use cases 3 and 4) there are differences between the models (Section 2.2.1). By additionally allowing for endogenous capacity expansion of flexibility options (use cases 2 and 4), complementary model differences to the case of exogenously given capacities (use cases 1 and 3) can have an effect. Finally, by considering different sets of VRE capacities and thus supply shares, it is possible to investigate to what extent model-related differences are correlated with this central parameter of energy supply. The exact design of these three scenario dimensions is explained in more detail in the following.

The reduced systems of use cases 1 and 2 include exogenous capacities of photovoltaics (PV), wind onshore, and wind offshore as VRE technologies, and battery storage, gas turbines, and transmission lines to balance them (Figure 2). The capacity optimization in use cases 2A-2D includes battery storage and gas turbines. Accordingly, no existing plants are assumed for these. The technology portfolio in use cases 3 and 4 includes numerous other technologies as shown in Figure 2. Additional dispatchable generation, electricity storage, demand response (DR) and flexible sector coupling provide the system with additional flexibility to balance VRE power generation. In use cases 4A-4D, the capacities of these flexibility options are also partially optimized. Differences in the technology scope of the models in use cases 3 and 4, as well as their compensation, are presented in Section 2.2.1.

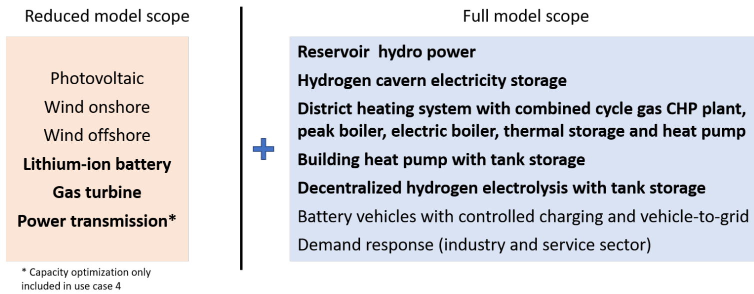

Figure 2: Specification of the model scope in the uses cases. While the use cases 1 and 2 (left side) consider only a small number of technologies, use cases 3 and 4 (right side) include numerous additional flexibility options. Bold technologies are available for capacity expansion in use case 2 and 4 in a subset of models according to Figure 3.

Across all use cases, a stylized system with 11 model nodes is considered. This corresponds approximately to the countries of Central Europe (Austria, Belgium, Denmark, Czech Republic, France, Germany, Italy, Luxembourg, The Netherlands, Poland, Switzerland) in terms of electricity demand, VRE potentials and time profiles. As for sector coupling, as well as the potentials of reservoir hydro power and hydrogen cavern storage, stylized assumptions are made that do not necessarily reflect real-world conditions.

The exogenously assumed VRE capacities are designed to be theoretically sufficient to supply 40% (A), 80% (B), 120% (C), and 160% (D) of demand in the reduced system (use cases 1 and 2). These theoretical shares may be reduced due to VRE curtailment and losses. By assuming higher demand when considering sector coupling and the full model scope (use cases 3 and 4), the theoretical VRE shares are reduced to 28% (A), 57% (B), 85% (C), and 114% (D). The annual electricity demand amounts to 3020 TWh in the case of the reduced system consideration and to 4240 TWh in the case of the full system consideration. Which supply shares can actually be realized depends on the availability and operation of the flexibility options and is analyzed in Section 3.1.

The exogenously defined capacities of electricity storage and dispatchable power plants are sized to the residual peak load occurring in each country, i.e., the maximum value of the difference between demand and VRE generation. These capacities are adopted here for battery storage, hydrogen cavern storage, gas-fired power plants, and combined heat and power (CHP) plants. By this assumption, there is a structural oversupply of flexibility in use cases 1 and 3, as there is at least twice the maximum required capacity. In order not to exacerbate this further, the consideration of reservoir hydro power plants is limited to a selection of model regions (Austria, Switzerland, Czech Republic, France, and Italy). The capacities of the technologies for flexible sector coupling (HPs, BEVs, and hydrogen electrolysers) were designed in [7] for a uniform electricity demand and are adopted directly. For the power grid, the existing capacities according to [27] are assumed, and for DR the potentials according to [28]. The assumption of a carbon dioxide (CO_2) price of 107€/t is a strong driver for the use of the different flexibility options considered. All model input data are available at (link)[1].

2.2. Contributing models

The model comparison includes a total of eight established power system models with sector coupling. These are DIETER [29, 30], E2M2 [31, 32], GENESYS-2 [33, 34], ISAaR [35, 36, 37], JMM [38], MarS [39], oemof [40, 41, 42] and REMix [43, 2]. A detailed overview of the fundamental model characteristics is provided in [7], where an earlier comparison of the models with a different focus is presented.

All models minimize the cost of technology operation over the course of a year at hourly resolution. To complement this, the use cases 2 and 4 also consider the proportionate investment cost of endogenous capacity installations. In the configuration used here, all models are formulated as linear, non-integer problems. Since the models use a harmonized data set, differences in the results may arise for three reasons: different technology scope (Section 2.2.1), different modeling approaches (Section 2.2.2), and different technology modeling (Section 2.2.3).

2.2.1. Differences in model scope

While the models are fully harmonized in their technology scope in the use cases 1 and 2, there are differences in the considered technologies in the use cases 3 and 4 (Figure 3). Capacity optimization is only relevant for cases 2 and 4. These were not considered with the JMM and MarS models. The differences in model scope result partly from whether modeling of a technology is possible at all, partly from the trade-off between model complexity and solution time, and partly from model-specific choices in this model comparison.

The overview shows that there is no model pair with the same technology scope in the use cases 3 and 4. Thus, not only do differences in modeling approach and technology modeling interact, but also those in model scope. In order to derive comparable results, substitute technologies are partially taken into account for technologies not considered (Figure 4).

In the models that do not explicitly consider specific sector coupling technologies, such as BEVs, these are represented in a stylized way by including a respective additional, inflexible electricity demand. This ensures that the same electricity demand has to be met in all models. However, differences in electricity demand may arise if electricity is used for heat generation in district heating (DH) systems. In addition, there is a lower energy demand in the models without representation of CHP, since the heat demand of the corresponding

[1] The model input data is being prepared for a full open access publication on Zenodo. Upon acceptance, the link will be added here. The submitted material includes exemplary input files for use case 4

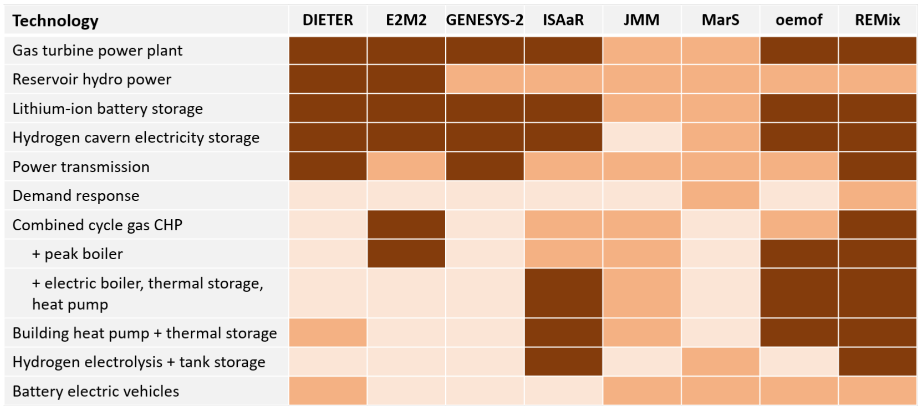

Figure 3: Technology scope of the models applied in the comparison. The dark color indicates an endogenous capacity expansion, the middle color a consideration of exogenous capacities and the light color a disregard of the corresponding technology. Please note that this overview does not necessarily reflect the general ability of the underlying models to consider these technologies.

DH systems is not considered elsewhere. This mostly affects the reported systems costs, which do not include the provision of this heat. To consider equal capacities of controllable power plants, CHP plants are replaced by additional gas-fired power plants if they are not considered.

2.2.2. Differences in modeling approach

With the exception of GENESYS-2 and JMM, all models optimize with perfect foresight over the overall time horizon of one year. JMM uses a rolling planning horizon for the optimization of the yearly dispatch. The year is divided into smaller periods of one week that can be solved successively to lower the complexity of the overall problem. In contrast to all other models, GENESYS-2 does not rely on a deterministic optimization, but on a population-based heuristic. Furthermore, it is designed as a dispatch model with every time step being solved independently without any foresight. In doing so, the use of technologies follows a predefined order. This order prefers local use or storage of energy before transmission. Only if there is a local surplus of VRE generation, transmission is considered. Then, the region under consideration tries to distribute this surplus starting with all neighboring regions and only going beyond them to more distant regions if necessary. Eventually, this leads to a more regional use of VRE. If, after the distribution of VRE, there is a shortage in demand in one region, it can request generation of other power plants from the neighbors first and then beyond.

2.2.3. Differences in technology modeling

The models used have numerous minor and major differences in technology modeling. The relevant ones for the following analysis include:

Technology / Model scope	With capacity optimization	Dispatch optimization only	Not modelled
Gas turbine power plant	Green field expansion	Exogenous capacities	Not applicable
Reservoir hydro power	Green field expansion	Exogenous capacities	Not applicable
Lithium-ion battery storage	Green field expansion	Exogenous capacities	Not applicable
Hydrogen cavern electricity storage	Green field expansion	Exogenous capacities	No consideration
Power transmission	Green field expansion	Exogenous capacities	Not applicable
Demand response	Not applicable	Exogenous capacities	No consideration
Combined cycle gas CHP	Green field expansion	Exogenous capacities	Doubled gas turbine capacity (CapFix), No consideration (CapOpt)
+ peak boiler	Green field expansion	Exogenous capacities	Not applicable
+ electric boiler, thermal storage, heat pump	Green field expansion	Exogenous capacities	Only CHP and boiler considered
Building heat pump + thermal storage	Green field expansion	Exogenous capacities	Exogenous power demand profile
Hydrogen electrolysis + tank storage	Green field expansion	Exogenous capacities	Exogenous power demand profile
Battery electric vehicles	Not applicable	Exogenous capacities	Exogenous power demand profile

Figure 4: Strategy for indirectly considering technologies in the models where they are not explicitly modeled. 'Not applicable' indicates that a certain configuration is not present in any of the models, 'No consideration' implies that no substitute for a technology is considered in the models.

Power plant outages. Power plants outages have been modeled using two different approaches. In most models, outages are interpreted as a certain percentage of continuous unavailable generation capacity. This implicitly assumes an equal distribution of outages over all hours of the year. A different approach is the stochastic drawing of outages which results in a partial or full unavailability of power plants, which is applied in MarS. While ISAaR considers outages only for exogenous capacities, oemof does not consider outages at all, which implies that all assets are available with their nominal capacity in each hour of the year.

Power plant ramping. In JMM and ISAaR, start up processes of power plants are associated with an additional fuel consumption. This leads to a higher overall fuel demand and therefore to an increase in CO_2 emissions.

Representation of reservoir hydro power plants. A simplified representation of hydro power plants, where hydraulic plants are modeled as an aggregated single unit, has been used by most of the models. In GENESYS-2, natural inflows are neglected. In MarS, a cascading model is implemented which allows that natural inflows are used multiple times. In JMM, the use of hydro reservoirs is determined by a water value which is calculated model-endogenously based on a reference electricity price and a reference filling level.

Storage and reservoir expansion. E2M2 and ISAaR use exogenously defined energy to power (E2P) ratios for some of the endogenously built storage technologies. In all other models this ratio is optimized. E2M2 applies the exogenously defined E2P ratios considered in use case 3 to the endogenously built capacities in use case 4. For the electric energy storage units, an E2P ratio of 4 h for batteries and 400 h for hydrogen caverns is assumed. For reservoir hydro power an E2P ratio of 615 h for the pump and 400 h for the turbine is considered. In ISAaR a fixed E2P ratio of 10.4 h is assumed for thermal energy storage (TES) expansion. Furthermore, in E2M2 the charging and discharging capacities of electricity storage must be identical, just as in ISAaR for TES. In the other models, this is only required for battery storage.

Power transmission. The most relevant difference in the representation of power transmission is the consideration of a direct current (DC) load flow approach in REMix. In contrast, all other models consider a net transfer capacity (NTC) approach, which allows higher line utilization to be realized. Transmission losses are considered in all models except MarS.

Battery electric vehicles. Differences in the modeling of BEVs particularly concern the calculation of costs. No costs are incurred for controlled charging in JMM, MarS and oemof. In DIETER, the same specific costs are applied for each charging process independent of the timing. In REMix, costs are only incurred if there is a deviation from the exogenously specified profile of uncontrolled charging. Costs for feeding electricity back into the grid are applied in all models except MarS. The modeling in JMM also differs from the other models in that no minimum battery level is considered. In addition, it is assumed that vehicles are always fully charged before driving and are reconnected to the grid with a predefined battery level.

Thermal and hydrogen storage. There is no bypass available for building HP storage in DIETER and JMM. The same is true for hydrogen tank storage in MarS. This implies that the entire production must pass through the storage system. From this follows that the reported amounts of stored energy are larger and higher losses can occur.

CHP fuel costs. In E2M2, CHP fuel consumption is based on the equivalent electricity generation. This is calculated as the sum of electricity and heat generation, but the latter is multiplied by the power loss factor and represents the equivalent electricity generation at which the same amount of fuel is consumed for the generation of pure electricity as for the actual combined generation of electricity and heat [31].

2.3. Output indicators

The evaluation of the model comparison is essentially based on the comparison of central parameters of the technology operation. Thus, annual values of energy provision, VRE curtailment, unsupplied demand, and system costs are compared for the overall assessment area. When endogenous plant installations are considered, capacities are additionally evaluated. We use normalized indicators to allow for better comparison of outcome deviations. In use cases 1 and 3, the reported system costs represent the variable operating costs of all assets including fuel and CO_2 costs. In use cases 2 and 4, these also include the annuities of investment costs of endogenously added assets and their fixed operating costs, but no investment costs for exogenous capacities. In use cases 3 and 4, unsupplied demand can include heat and hydrogen in addition to electricity.

3. Results and discussion

The evaluation and analysis of the results starts with the key indicator for the usage of flexibility options, which is the realized VRE share (Section 3.1). It then follows the structure of the model comparison shown in Figure 1. Thus, the use cases with reduced, harmonized model scope without (Section 3.2) and with capacity optimization (Section 3.3) are considered first, and then those with full, heterogeneous model scope, also first without (Section 3.4) and then with capacity optimization (Section 3.5). Finally, the relation between VRE technology share and deviations in model outcomes is evaluated (Section 3.6).

3.1. Realized VRE shares

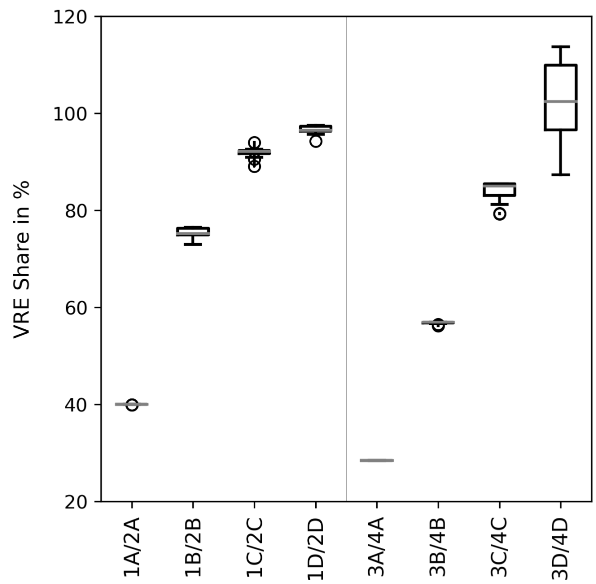

Figure 5: Range of realized VRE supply shares in the systems with reduced model scope (use cases 1 and 2, left side) and full model scope (use case 3 and 4, right side).

Despite the numerous model differences, we find a high robustness of the VRE supply share calculated based on the used wind and PV power generation and the exogenous electricity demand (Figure 5). In general, the deviations increase with rising VRE capacity and are higher in the cases with full model scope than with reduced model scope. In the use cases with lowest VRE capacity (A), differences of less than 0.1% arise because the VRE electricity generation can be fully utilized. With reduced model scope (use case 1 and 2), higher VRE capacities trigger differences in the realized VRE shares of up to about 5%. A much larger spread results in the case of the full model scope and the highest VRE capacities (D). Due to the different consideration of flexibility options, the difference in the achieved VRE shares reaches up to 25% there. In particular, the use of VRE for heat generation in DH contributes to this, which also enables VRE supply shares of more than 100%. Considering an endogenous capacity expansion leads to lower VRE supply shares for both reduced and full model scope. However, the differences in results caused by endogenous expansion are significantly smaller than those between the models. Due to storage and grid losses, the VRE share in the final electricity supply may be lower and show larger differences between the models than the values shown here. This is analyzed in detail in the following.

3.2. Reduced model scope without capacity expansion (use case 1)

Figure 6 shows the key indicators for use cases 1A-1D. There, all models consider only gas turbines, battery storage, and power transmission as flexibility options.

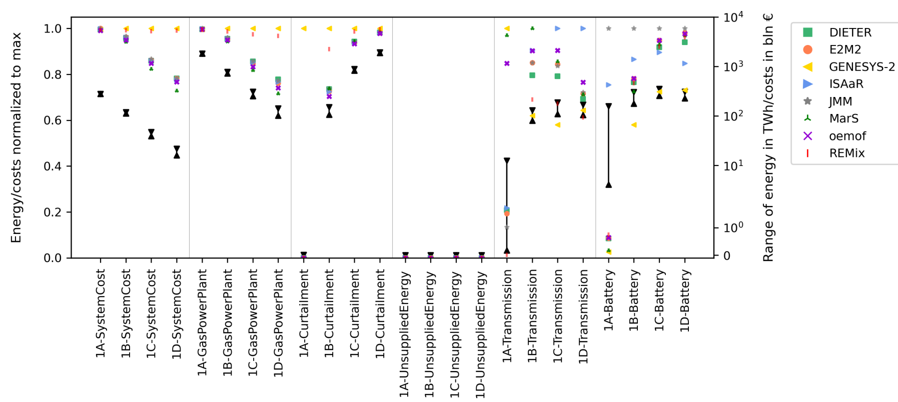

Figure 6: Key power system operation indicators in use case 1, including system costs, power generation in gas power plants, VRE curtailment, unsupplied energy, power transmission, and battery storage output. The colored symbols show model-specific values for each use case and indicator normalized to the corresponding maximum according to the scale on the left y-axis. The black ranges indicate the absolute values on the logarithmic scale of the right y-axis.

The results reveal several deviations between the models and some clear trends. All models find a strong decrease in gas power plant usage and system costs (variable operating costs) with increasing VRE supply share, while renewable curtailment increases. The use of power storage and transmission first increases but then reaches a level of saturation at very high VRE shares when excess production occurs more frequently both in time and space. In addition, we find that model-related differences in system costs and power plant dispatch increases with VRE share, whereas it decreases for grid usage, storage usage, and VRE curtailment.

The most significant deviations in model results can be clearly associated to the model differences. The largest differences are related to the use of a DC load flow approach and the assumption of no foresight over time, smaller ones due to fixed dispatch order, power plant ramping, power plant outages, and grid losses.

For the system costs, there is a very high agreement in the results across most models. Outliers arise for higher VRE shares due to deviating grid consideration (REMix) and a dispatch approach without temporal foresight (GENESYS-2). Since the system costs are essentially driven by the variable costs of gas-fired power plant operation, a very similar pattern is observed for the latter. There, the deviation is about 25% for the outliers, and consistently less than 10% for the other models. The lowest power plant utilization and thus also the lowest costs result from neglecting the grid losses (MarS).

A structurally analogous pattern also emerges in the case of the VRE curtailment, which deviates from one another in most models only in the single-digit percentage range, with the aforementioned outliers of REMix and GENESYS-2. A more heterogeneous picture is found for the use of the electricity grid and battery storage. In both cases, very large relative deviations of over 80% arise in the case of low VRE shares (1A), which correspond to only small absolute deviations due to the low use of the technologies. In the case of higher VRE shares, the deviation then reduces again to values below 50%, excluding the outliers even to below 20% (power grid) and 10% (battery storage). With respect to grid usage, outliers can be explained by the modeling differences. Stochastic power plant outages are partially compensated by energy imports, which can lead to a substantial relative increase in grid usage at low VRE shares (MarS).

Using the DC load flow approach reduces the available grid capacity, which translates into lower grid utilization, which in turn must be compensated by gas-fired power plants (REMix). The application of a predefined dispatch order makes the whole system more inflexible and inefficient (GENESYS-2). This results in higher system cost, generation from gas power plants, and curtailment, while at the same time reducing battery and transmission usage. Furthermore, a more detailed modeling of gas power plants involving additional higher fuel consumption for start-up processes results in a higher battery storage usage to smooth the gas power plant generation and therefore to reduce power plant start-ups (JMM, ISAaR). This effect can especially be observed at lower VRE shares where the power plant production is comparatively high.

3.3. Reduced model scope with capacity expansion (use case 2)

Use cases 2A-2D also consider a reduced but uniform technology scope, but with model endogenous optimization of battery storage and gas-fired power plant capacities. This deviation in model configuration does not change the trends in system operating parameters observed in the use cases 1A-1D for the increasing VRE supply share (Figure 7). Thus, we see an increase in curtailment, a decrease in power plant dispatch and system costs, and an initial increase and later saturation in the use of battery storage and transmission lines. These trends emerge equally for endogenous investments in battery storage and gas-fired power plants (Figure 8). In contrast to use case 1, the spread of model results increases for higher VRE shares for all indicators except for

VRE curtailment. This seems plausible considering the increasing endogenous plant expansion and its differences between the models.

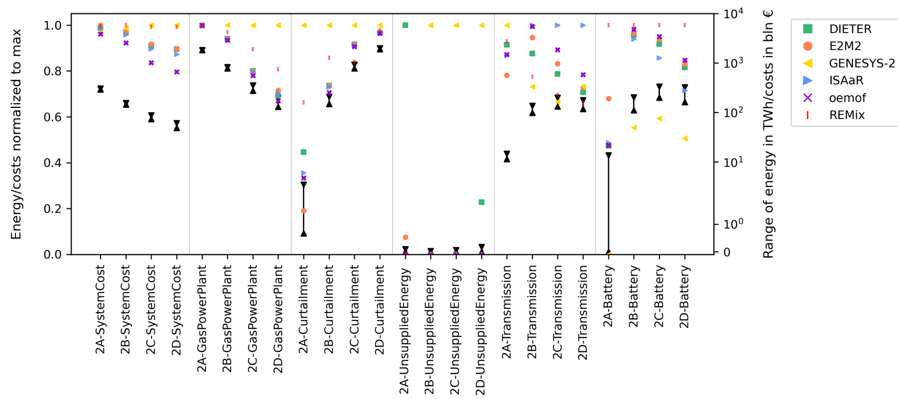

Figure 7: Key power system operation indicators in use case 2, including system costs, power generation in gas power plants, VRE curtailment, unsupplied energy, power transmission, and battery storage output. The colored symbols show model-specific values for each use case and indicator normalized to the corresponding maximum according to the scale on the left y-axis. The black ranges indicate the absolute values on the logarithmic scale of the right y-axis.

In the endogenous addition of battery storage and gas-fired power plants, there is a high agreement between most of the models (Figure 8). However, there are also significant differences in the preferred technology and the total capacity added, which can be explained by the model differences. For example, the separate optimization of the individual time steps strongly favors the addition of gas-fired power plants, which is why battery storage is added to a lesser extent (GENESYS-2). In contrast, when the exogenously defined E2P is lower than the optimal value, larger battery converter capacities must be provided to obtain a similar energy storage capacity as most other models (E2M2). Less pronounced is the impact of a lower usable power transmission capacity, which results in higher capacities for both storage and power plants (REMix). If a full availability of generators is considered, lower plant capacities are systematically required (oemof, ISAaR). Due to these model differences, the aggregated power generation capacity of endogenously added battery storage and gas-fired power plants differs by a maximum of 30%.

The system operation parameters show essentially the same characteristics as in the systems without endogenous capacity expansion (cf. Figure 6 and Figure 7). This applies equally to the trends across the models and to the deviations between the models. For example, model results are relatively similar in terms of system costs and gas-fired power plant operation. Deviations between the models increase with the VRE share and reach a maximum of 10% in use case 2D. Outliers upwards are again the models with deviating grid modeling (REMix) and modeling methods (GENESYS-2). In contrast, a downward outlier in the costs results from the neglect of plant availability, which reduces the

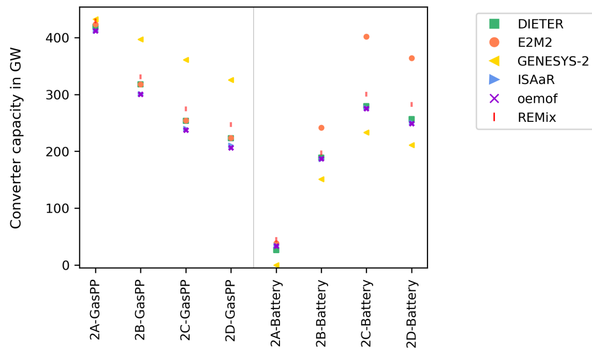

Figure 8: Endogenous power generation capacity installations for gas power plants and battery storage converters in use case 2. In use case 1, the exogenous capacities of both technologies amount to 492 GW.

specific investment costs (oemof). An inflexible predefined dispatch order can cause small amounts of unsupplied energy in this use case (GENESYS-2). It is noticeable that in the case of endogenous expansion, compensation for the lower grid capacities is made to a much greater extent by batteries, which can reduce the use of gas-fired power plants compared to the case with exogenous capacities (REMix).

3.4. Full model scope without capacity expansion (use case 3)

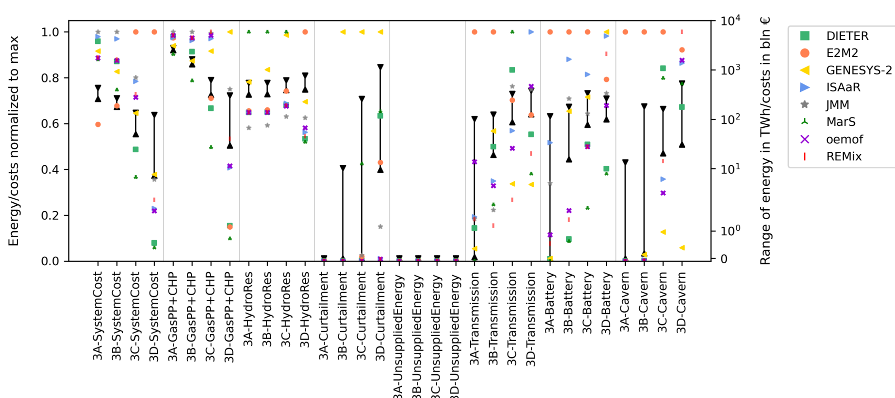

Figure 9: Key power system operation indicators in use case 3, including system costs, power generation in gas-fired power as well as CHP plants, reservoir hydro power plants, VRE curtailment, unsupplied energy, power transmission, battery storage output, and cavern storage output. The colored symbols show model-specific values for each use case and indicator normalized to the corresponding maximum according to the scale on the left y-axis. The black ranges indicate the absolute values on the logarithmic scale of the right y-axis.

In the use cases 3A-3D, additional technologies are added to the power system to balance VRE power generation, as shown in Figure 2. In doing so, all models consider only part of the full technology spectrum (Figure 3). Nevertheless, the expected trends emerge for the main power system indicators (Figure 9). While the costs and utilization of gas-fired power plants, here including CHP plants, decrease with the VRE share, the utilization of storage and the power grid, as well as VRE curtailment, increase. The relative deviations between the models also show different trends. While they increase with the VRE share for the costs and use of gas-fired power plants, opposite trends emerge for transmission and storage. The differences for reservoir hydro power plants are relatively constant. Remarkably, there are large differences for VRE curtailment, which can be avoided in some models across all use cases, but reaches significant amounts in others. These substantially larger discrepancies between the results compared to the harmonized use cases can be explained by the identified model differences.

For example, neglecting electrical heat production in DH leads to significantly higher costs at high VRE shares, as more fuel is needed in conventional boilers (E2M2). In contrast, significantly lower system costs result if DH is not modeled, since the corresponding heat demand is not considered in the models and thus lower fuel costs are incurred (DIETER, MarS). Lower fuel costs can also be associated to the consideration of an equivalent electricity generation for calculating fuel consumption in CHP plants (E2M2).

The increasingly divergent power generation in gas-fired power and CHP plants with higher VRE shares is closely correlated with the available other flexibility. For example, a more detailed consideration of reservoir hydro plants can increase their electricity generation at the expense of gas-fired plants with rising VRE shares until saturation where the additional available energy cannot be integrated into the system (MarS). On the other hand, reduced storage possibilities over the long term due to limited time foresight causes higher use of thermal power plants (JMM, GENESYS-2).

The use of electricity for heat generation is clearly reflected in a reduction of VRE curtailment. For this, the long-term storage is also of great relevance, which is used a lot especially where electric heat generation is not considered (E2M2). More intensive use of long-term storage is accompanied by a decline in electricity generation from gas-fired power plants and CHP. Where both electric heating and long-term storage are available, no curtailment is observed even with high VRE shares (ISAaR, oemof, REMix). Instead, significantly higher VRE curtailment results if flexible sector coupling is not available and long-term storage is not operated optimally due to a lack of temporal foresight (GENESYS-2).

The use of the power grid shows few consistent model-specific trends. Analogous to use cases 1 and 2, the use of a DC load flow approach reduces the amount of electricity transmitted (REMix). Furthermore, the increased range of flexible technologies diminishes the impact of a stochastic modeling of outages, which resulted in increased grid usage in use case 1 (MarS).

Using a perfect foresight approach, the lower availability of flexible sector coupling options can be partially compensated by a more intensive use of the power grid, reservoir hydro power, and storage (E2M2). In contrast, cavern storage is not used at all in the other models when VRE shares are below 90% (use case 3C), and battery storage is used almost exclusively to reduce power plant startups (ISAaR, JMM).

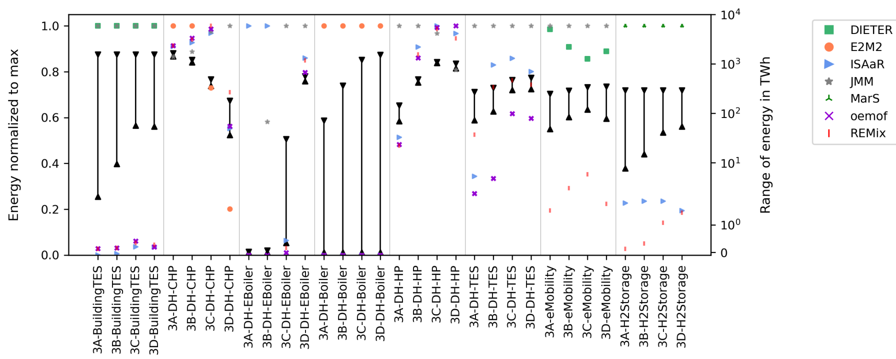

Figure 10: Sector coupling operation indicators in use case 3, including heat production and storage, load shifting of BEVs, and hydrogen tank storage usage. The colored symbols show model-specific values for each use case and indicator normalized to the corresponding maximum according to the scale on the left y-axis. The black ranges indicate the absolute values on the logarithmic scale of the right y-axis.

The operation of flexible sector coupling technologies (Figure 10) also shows clear trends. CHP heat is replaced with increasing VRE share, either by a conventional peak load boiler (E2M2), or, where possible, by electric heat generation with HP and electric boiler (ISAaR, JMM, oemof, REMix). In the case of electric heat generation, TES is used in a complementary way for flexibility. The flexibility of hydrogen electrolysis is also used more as the VRE share increases, whereas it does not show a clear trend for BEVs.

Even though the electrification and flexibilization of the DH supply is clearly visible in all models, there are significant differences in implementation. This applies in particular to the use of TES, where differences of more than 70% arise. When hydrogen cavern storage is not available, TES serves as an alternative option for long-term balancing (JMM). This is accompanied by more intensive use of electric heat generation. Moreover, a higher use of BEV flexibility can be observed.

The large differences in the use of building TES and hydrogen tank storage result from the fact that these do not have a bypass in some models (DIETER, JMM, MarS), which causes the reported discharging to be significantly higher. Nevertheless, different flexibility requirements lead to significant differences in storage usage even in comparable models. For BEVs, different technology modeling approaches cause significant variations in the flexibility provided. In particular, the application of costs to a deviation from a predefined charging profile has a significant reducing effect on the use of BEV flexibility (REMix). In contrast, neglecting a minimum vehicle battery level increases the use of flexibility

(JMM). If flexible charging of BEVs is not possible, other flexibility options such as battery storage and hydrogen storage are used more extensively (E2M2, ISAaR).

3.5. Full model scope with capacity expansion (use case 4)

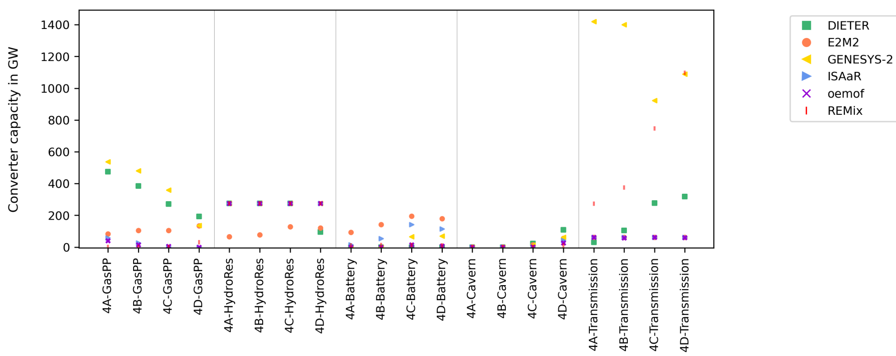

Figure 11: Available capacities of power plants, storage, and transmission lines in use case 4. Exogenous capacities are shown here for some models in reservoir hydro power (HydroRes) and power lines (Transmission), as can be seen from the identical values in all use cases and Figure 3.

In use cases 4A-4D, the full system analysis also includes endogenous capacity expansion. This gives the models numerous additional degrees of freedom. These are used in different ways, resulting in even greater deviations of results. This can be seen, for example, in the endogenous investment decisions. It should be noted that each model has different technologies available for an expansion (Figure 3). This results in large deviations in investment decisions for power plants, electricity storage, and power lines (Figure 11).

The significant differences in the expansion of gas-fired power plants can be explained by the consideration of CHP. Where CHP is not available, significantly more gas-fired power plants are needed (DIETER, GENESYS-2). In contrast, the consideration of exogenous CHP capacities makes gas power plants almost or even completely obsolete (ISAaR, oemof). In the case of endogenous CHP expansion, both technologies are used (E2M2). With a combined expansion of CHP and power transmission lines, gas-fired power plants can be almost completely avoided (REMix). If grid expansion is also possible, the option of expanding reservoir hydro power is used to the maximum as long as VRE generation does not exceed demand (DIETER). In contrast, if reservoir hydro power is optimized in capacity with predefined ratios and without grid expansion, lower capacities are built, which must be compensated by additional gas-fired power plants and batteries (E2M2). Beyond that, battery storage is only added endogenously to compensate for the reduced flexibility of CHP plants if additional start-up and load change costs are considered for these (ISAaR). A small endogenous expansion of hydrogen cavern storage occurs only at higher VRE shares and when it competes with reservoir hydro power (DIETER). An expansion of the power grid is realized where it is possible, generally increasing with the VRE share (DIETER, REMix). The use of a DC load flow approach leads to higher values for the absolute capacities (REMix). Allowing for grid expansion with a fixed dispatch order can lead to very high endogenous capacities at low VRE shares (GENESYS-2). This results from the fact that grid expansion is more attractive in the short term compared to reduction of VRE curtailment by the addition of further storage, as there is still unmet demand in other regions. Models with deterministic optimization can balance those events using other storage when VRE shares are low but with increasing VRE share they also require higher transmission capacity. In the case of non-deterministic optimization without the option of foresight (GENESYS-2), the flexibility of storage is therefore lower. However, the more balancing options are available the lower the required capacity expansion for transmission (oemof, REMix).

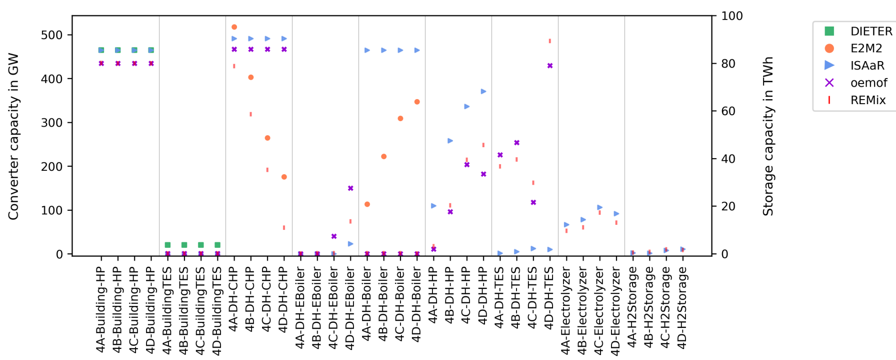

Figure 12: Available capacities of sector coupling flexibility options in use case 4. This includes converters in the heat and hydrogen supply (left axis) and storage energy capacities for thermal and decentralized hydrogen storage (right axis). Exogenous capacities are shown here for some models according to Figure 3.

There are also large discrepancies between the models in terms of investments in flexible sector coupling technologies (Figure 12). Where endogenous dimensioning of CHP is possible, it is chosen to be lower with increasing VRE share (E2M2, REMix). On the heat side, this is compensated either by additional gas boilers (E2M2) or by electric heat generators (REMix). However, an electrification of the DH supply is also made possible by appropriate investments if the CHP capacities are exogenously specified (ISAaR, oemof). With higher VRE shares, TES is increasingly being built for additional flexibility of DH supply (oemof, REMix). Considering a fixed E2P ratio for TES in DH networks reduces their optimal size, which in turn leads to higher HP capacities in these networks compared to the other models (ISAaR).

Decentralized TES proves to be too expensive, which is why no (ISAaR) or only a very minor expansion (oemof, REMix) takes place. With about 130 GWh, it is significantly lower than the exogenous capacities (see values for DIETER).

Due to the lack of flexibility, the endogenous capacities for the associated HPs are identical in all use cases. Depending on whether the coefficient of performance of these HPs is time-variable (oemof, REMix) or constant (ISAaR), the optimal capacities are slightly different. In the latter case, higher values result, which are identical to the exogenous capacities.

In the endogenous expansion of decentralized electrolysers and hydrogen storage, the two models involved show similar trends. Thus, the optimal electrolyzer capacities are consistently lower than the exogenously defined ones in use case 3 (108 GW). Moreover, an increase with rising VRE share is observed in 4A-4C, as well as a slight decrease in 4D. At below 2 TWh, aggregate hydrogen tank storage capacity remains at a very low level. Thereby, an increasing trend with the VRE share can be seen in both models.

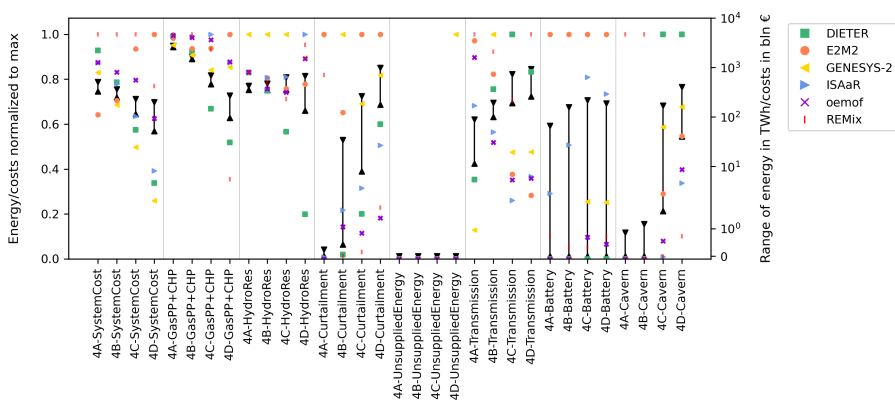

Figure 13: Key power system operation indicators in use case 4, including system costs, power generation in gas power as well as CHP plants, reservoir hydro power plants, VRE curtailment, unsupplied energy, power transmission, battery storage output, and cavern storage output. The colored symbols show model-specific values for each use case and indicator normalized to the corresponding maximum according to the scale on the left y-axis. The black ranges indicate the absolute values on the logarithmic scale of the right y-axis.

With respect to system operation, there are deviations of at least one order of magnitude between the models for most indicators and use cases (Figure 13). Apart from a few outliers, however, the trends in the dependencies between VRE share and technology deployment of the previous use cases are confirmed. At least the larger deviations between the results can again be explained by the model differences.

The strong upward outlier in system costs results from the consideration of an endogenous capacity expansion of all conversion plants and storage facilities of sector coupling (REMix). The associated investment costs do not apply if, as a substitute, only the electricity demand of sector coupling is taken into account in other models (Figure 4). In contrast, lower costs result where comparatively few technologies are optimized endogenously or the fuel requirement for supplying the heating networks is omitted (DIETER, GENESYS-2, ISAaR).

Regarding the lower costs at low VRE shares caused by consideration of the equivalent electricity generation of the CHP plants, as well as the increased costs at high VRE shares caused by the fuel demand of the peak load boilers, the same effects result as in use case 3 (E2M2). When using the gas-fired power plants, a difference arises in particular due to the endogenous grid expansion in case 4D (DIETER, REMix). For reservoir hydro power plants, the most relevant effect is that a lower endogenous capacity deployment substantially reduces their electricity supply in case 4D (DIETER, E2M2). From the need to invest in the various flexibility options follows that they are available in total at a significantly lower capacity. This goes hand in hand with a higher VRE curtailment (DIETER, E2M2, ISAaR, REMix). As expected, endogenous power grid expansion leads to significantly greater grid usage (DIETER, GENESYS-2, REMix). The significantly lower use of electricity storage compared to use case 3 is directly related to the low endogenous addition of these facilities.

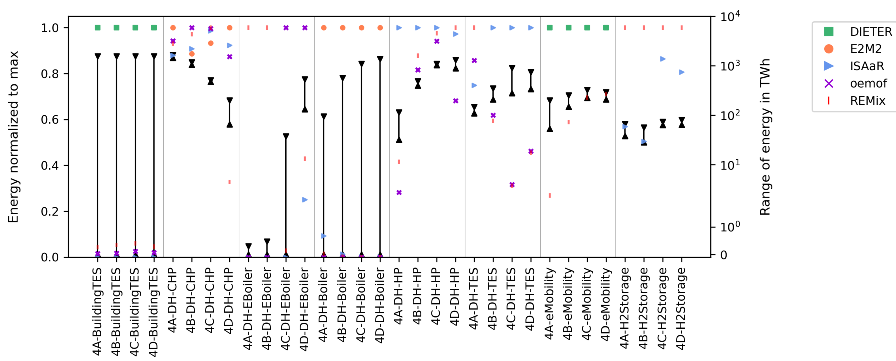

Figure 14: Sector coupling operation indicators in use case 4, including heat production and storage, load shifting of BEVs, and hydrogen tank storage usage. The colored symbols show model-specific values for each use case and indicator normalized to the corresponding maximum according to the scale on the left y-axis. The black ranges indicate the absolute values on the logarithmic scale of the right y-axis.

Considering an endogenous plant expansion does not result in substantial changes to the deployment of sector coupling technologies and its dependence on the VRE share (Figure 14). Model-specific effects can also be attributed to the same causes.

3.6. Regional effects

The consideration of a multi-node system allows the analysis of the dependence between model deviations and VRE supply share (Figure 15). This is examined using the VRE curtailment as an example. The average deviation from the median over all models is considered. Since there are numerous scenarios and models in use case 3 and 4 in which no VRE curtailment occurs, the analysis only includes use cases 1 and 2.

The results show mean deviations from the median of a maximum of 20%. There is a clear correlation with the distribution between wind and solar energy.

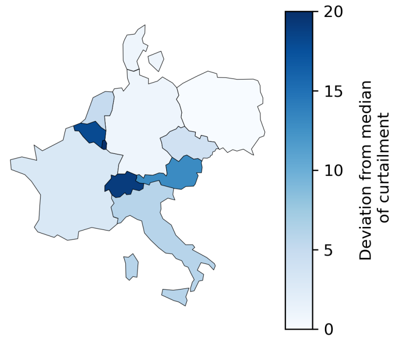

Figure 15: Regional differences in model deviations. Shown are the averaged deviations from the median of the VRE curtailment, averaged over all use cases 1 and 2.

While the median deviation in Poland, where 77% of the VRE electricity generation comes from wind energy and only 23% from PV, is 0%, the highest values of more than 15% are found in Belgium, Luxembourg and Switzerland, where PV accounts for at least 60% of the generation in each case. This suggests that the present model differences have a larger effect at high PV shares.

4. Conclusions

In a structured comparison, we quantify model-related deviations in the outcomes of eight power system models with sector coupling for 16 use cases. Our model results are very similar in case of a completely harmonized and reduced model scope with pure dispatch optimization of VRE, electricity storage, power transmission, and gas-fired power plants (use case 1). Here, the considered indicators of technology operation mostly deviate by less than 40%. For most of the models, the differences are even less than 20%. These general findings do not change if endogenous capacity optimization of gas-fired power plants and storage is also considered (use case 2). In general, the relatively high agreement of the results suggests a high robustness of the models, and can be taken as a sign of their validity.

As expected, the consideration of different technology scopes and sector coupling leads to significantly larger deviations in results than is the case with fully harmonized models. Even with complete specification of exogenous plant capacities (use case 3), there are deviations in technology operation of more than one order of magnitude in some cases, and even higher for some outliers. However, there are also subgroups of models with similar results, especially when the model scope differs only slightly. The ranges of results are even higher when additional model degrees of freedom are available by considering endogenous plant expansions of further flexibility options (use case 4). Nevertheless, it is possible to compare the results in a meaningful way and to explain the observed deviations by the identified model differences. While the deviations in the use of the individual flexibility options converge with reduced cases and exogenous capacities as the VRE share increases, the opposite picture emerges with a broad technology portfolio and capacity optimization.

Our analysis shows that different model scopes and modeling approaches have substantially larger impacts than the identified differences in technology modeling. With regard to the modeling approach, this relates here in particular to the use of a heuristic approach with a predefined technology dispatch order, and to a lower extent to the use of a rolling time horizon. In terms of technology scope, neglecting flexible supply of district heating with CHP and heat pumps as well as flexible charging of battery vehicles have a substantial impact on results. As regards technology modeling, differences for reservoir hydro power plants and the power grid lead to the largest deviations in plant deployment, whereas considering fixed storage designs and neglecting power plant outages lead to the largest deviations in investments.

Despite the sometimes very large deviations between the model results, robust effects emerge with regard to the use of technologies and their dependence on the VRE share. Consistent outcomes include, for example, the flexibility of sector coupling, which is used across all models to the extent that it is considered. This particularly concerns the flexibility of vehicle charging and the partially electric heat production in district heating. Moreover, we show that even with theoretical VRE shares well above 100% of demand and a broad portfolio of flexibility options, dispatchable power plants or CHP plants are not abandoned completely. The increased use of the power grid and long-term storage at higher shares of VRE is also consistent in all models. In contrast, there is no clear picture for the use of stationary batteries, as these are replaced by flexible sector coupling in some models. Nevertheless, whether and how individual flexibility options are used in the models is closely linked to the questions of which other technologies are considered and how they are modeled. This must be taken into account in the interpretation and evaluation of model results. The appearance of robust results in technology usage suggests that the models can be used to address the same issues despite their differences in detail. Still, the specific characteristics and specializations should be taken into account when selecting the model, as they can have a significant impact on the results.

With regard to the method used, our analysis shows that the use of harmonized input data and profound analysis of model properties allow the association of key outcome deviations with model differences. Thus, the effects of different modeling approaches can be captured and quantified. In addition, the effects of considering or not considering individual flexibility options on the operation of the modeled system can be analyzed. Most of the analyzed effects can be observed in several models. This suggests that these findings on the effect of model differences can also be applied to other models based on a cost minimization approach. This gives other modelers and users of model-based energy scenarios the possibility to better interpret the results. To further strengthen the understanding of model robustness, future comparison studies should consider larger sets with identical model scope even for a full analysis of all flexibility options. While models that only cover the traditional power sector are more easy to harmonize, comparative studies of models with numerous flexibility and sector coupling technologies are more challenging. This concerns the harmonization of both the technology scope and of cross-sectoral input data. Here we see a promising field for future research on more detailed model comparisons.

By considering different VRE shares and several regions in the scenarios, a broad spectrum of supply structures is taken into account. This increases the possibility of transferring general findings to more realistic scenarios. Nevertheless, follow-up studies should complementarily quantify the impact of model choices on outcomes of realistic transformation scenarios.

Acknowledgements

The research for this paper was performed within the FlexMex project supported by the German Federal Ministry of Economic Affairs and Energy under grant number 03ET4077. The authors thank Georgios Savvidis, Felix Böing, Kai Hufendiek, Benjamin Fleischer, Christoph Weber, Berit Müller, Mascha Richter, Frank Merten, Dirk-Uwe Sauer, Niklas van Bracht, Arne Pöstges, and Caroline Podewski for contributing to the project conceptualization and funding acquisition. Furthermore, we thank Alexander Fehler, Alexander Zerrahn, Christine Krüger and Tomke Janßen for their contribution to model preparation and discussion on the models.

Author contribution

Hans Christian Gils: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing- Original draft preparation, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition **Hedda Gardian:** Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing- Original draft preparation, Visualization **Martin Kittel:** Methodology, Software, Validation, Formal analysis, Investigation, Writing- Original draft preparation, Writing - Review & Editing, Visualization **Wolf-Peter Schill:** Investigation, Writing- Original draft preparation, Writing - Review & Editing, Funding acquisition **Alexander Murmann:** Methodology, Software, Validation, Formal analysis, Investigation, Writing- Original draft preparation, Visualization **Jann Launer:** Methodology, Software, Validation, Formal analysis, Investigation, Writing- Original draft preparation, Writing - Review & Editing **Felix Gaumnitz:** Methodology, Software, Validation, Formal analysis, Investigation, Writing- Original draft preparation, Writing - Review & Editing **Jonas van Ouwerkerk:** Methodology, Software, Validation, Formal analysis, Investigation, Writing- Original draft preparation, Writing - Review & Editing **Jennifer Mikurda:** Methodology, Software, Validation, Formal analysis, Investigation, Writing- Original draft preparation, Writing - Review & Editing **Laura**

Torralba-Díaz: Methodology, Software, Validation, Formal analysis, Investigation, Writing- Original draft preparation, Writing - Review & Editing

Data Availability

The input data and the data template used are available on [link][2].

References

[1] H. Blanco, A. Faaij, A review at the role of storage in energy systems with a focus on power to gas and long-term storage, Renewable and Sustainable Energy Reviews 81 (2018) 1049–1086. doi:https://doi.org/10.1016/j.rser.2017.07.062.
URL https://www.sciencedirect.com/science/article/pii/S1364032117311310

[2] H. C. Gils, H. Gardian, J. Schmugge, Interaction of hydrogen infrastructures with other sector coupling options towards a zero-emission energy system in germany, Renewable Energy 180 (2021) 140–156. doi:https://doi.org/10.1016/j.renene.2021.08.016.
URL https://www.sciencedirect.com/science/article/pii/S0960148121011769

[3] D. B. Richardson, Electric vehicles and the electric grid: A review of modeling approaches, impacts, and renewable energy integration, Renewable and Sustainable Energy Reviews 19 (2013) 247–254. doi:https://doi.org/10.1016/j.rser.2012.11.042.
URL https://www.sciencedirect.com/science/article/pii/S1364032112006557

[4] A. Bloess, W.-P. Schill, A. Zerrahn, Power-to-heat for renewable energy integration: A review of technologies, modeling approaches, and flexibility potentials, Applied Energy 212 (2018) 1611–1626. doi:https://doi.org/10.1016/j.apenergy.2017.12.073.
URL https://www.sciencedirect.com/science/article/pii/S0306261917317889

[5] E. S. Hanley, J. P. Deane, B. P. Ó Gallachóir, The role of hydrogen in low carbon energy futures – a review of existing perspectives, Renewable and Sustainable Energy Reviews 82 (2018) 3027–3045. doi:https://doi.org/10.1016/j.rser.2017.10.034.
URL https://www.sciencedirect.com/science/article/pii/S1364032117314089

[2]The model input data is being prepared for a full open access publication on Zenodo. Upon acceptance, the link will be added here. The submitted material includes exemplary input files for use case 4

[6] J. H. Kwakkel, W. E. Walker, V. A. Marchau, Classifying and communicating uncertainties in model-based policy analysis, International Journal of Technology, Policy and Management 10 (4) (2010) 299–315. doi:10.1504/IJTPM.2010.036918.
URL http://www.inderscience.com/link.php?id=36918

[7] H. C. Gils, H. Gardian, M. Kittel, W.-P. Schill, A. Zerrahn, A. Murmann, J. Launer, A. Fehler, F. Gaumnitz, J. van Ouwerkerk, C. Bußar, J. Mikurda, L. Torralba-Díaz, T. Janßen, C. Krüger, Modeling flexibility in energy systems - a scenario-based comparison of power sector models, Renewable and Sustainable Energy Reviews (under review) (2021).
URL https://elib.dlr.de/143072/

[8] Leonardo Gacitua, Pablo Gallegos, Rodrigo Henriquez Auba, Álvaro Lorca, Matías Negrete Pincetic, Daniel Olivares, A. Valenzuela, George Wenzel, A comprehensive review on expansion planning: Models and tools for energy policy analysis, Renewable and Sustainable Energy Reviews 98 (2018) 346–360. doi:https://doi.org/10.1016/j.rser.2018.08.043.
URL https://www.sciencedirect.com/science/article/pii/S1364032118306269

[9] Peter Lopion, Peter Markewitz, Martin Robinius, Detlef Stolten, A review of current challenges and trends in energy systems modeling, Renewable and Sustainable Energy Reviews 96 (2018) 156–166. doi:10.1016/j.rser.2018.07.045.
URL https://www.sciencedirect.com/science/article/pii/S1364032118305537

[10] Georgios Savvidis, Kais Siala, Christoph Weissbart, Lukas Schmidt, Frieder Borggrefe, Subhash Kumar, Karen Pittel, Reinhard Madlener, Kai Hufendiek, The gap between energy policy challenges and model capabilities, Energy Policy 125 (2019) 503–520. doi:10.1016/j.enpol.2018.10.033.

[11] Matteo Giacomo Prina, Giampaolo Manzolini, David Moser, Benedetto Nastasi, Wolfram Sparber, Classification and challenges of bottom-up energy system models - A review, Renewable and Sustainable Energy Reviews 129 (2020). doi:10.1016/j.rser.2020.109917.
URL http://www.sciencedirect.com/science/article/pii/S1364032120302082

[12] Amirhossein Fattahi, Jos Sijm, André Faaij, A systemic approach to analyze integrated energy system modeling tools: A review of national models, Renewable and Sustainable Energy Reviews 133 (2020). doi:10.1016/j.rser.2020.110195.
URL https://www.sciencedirect.com/science/article/pii/S1364032120304858

[13] Christian Klemm, Peter Vennemann, Modeling and optimization of multi-energy systems in mixed-use districts: A review of existing methods and approaches, Renewable and Sustainable Energy Reviews 135 (110206) (2021). doi:10.1016/j.rser.2020.110206.
URL http://www.sciencedirect.com/science/article/pii/S1364032120304950

[14] M. Sugiyama, S. Fujimori, K. Wada, S. Endo, Y. Fujii, R. Komiyama, E. Kato, A. Kurosawa, Y. Matsuo, K. Oshiro, F. Sano, H. Shiraki, Japan's long-term climate mitigation policy: Multi-model assessment and sectoral challenges, Energy 167 (2019) 1120–1131. doi:https://doi.org/10.1016/j.energy.2018.10.091.
URL https://www.sciencedirect.com/science/article/pii/S0360544218320814

[15] J. E. Bistline, M. Brown, S. A. Siddiqui, K. Vaillancourt, Electric sector impacts of renewable policy coordination: A multi-model study of the north american energy system, Energy Policy 145 (2020) 111707. doi:https://doi.org/10.1016/j.enpol.2020.111707.
URL https://www.sciencedirect.com/science/article/pii/S0301421520304341

[16] H. G. Huntington, A. Bhargava, D. Daniels, J. P. Weyant, C. Avraam, J. Bistline, J. A. Edmonds, S. Giarola, A. Hawkes, M. Hansen, P. Johnston, A. Molar-Cruz, M. Nadew, S. Siddiqui, K. Vaillancourt, N. Victor, Key findings from the core north american scenarios in the emf34 intermodel comparison, Energy Policy 144 (2020) 111599. doi:https://doi.org/10.1016/j.enpol.2020.111599.
URL https://www.sciencedirect.com/science/article/pii/S0301421520303372

[17] S. Giarola, A. Molar-Cruz, K. Vaillancourt, O. Bahn, L. Sarmiento, A. Hawkes, M. Brown, The role of energy storage in the uptake of renewable energy: A model comparison approach, Energy Policy 151 (2021) 112159. doi:https://doi.org/10.1016/j.enpol.2021.112159.
URL https://www.sciencedirect.com/science/article/pii/S0301421521000288

[18] Hans Christian Gils, Thomas Pregger, Franziska Flachsbarth, Mareike Jentsch, Constantin Dierstein, Comparison of spatially and temporally resolved energy system models with a focus on Germany's future power supply, Applied Energy 255 (2019). doi:10.1016/j.apenergy.2019.113889.

[19] Kais Siala, Mathias Mier, Lukas Schmidt, Laura Torralba Díaz, Siamak Sheykkha, Georgios Savvidis, Which model features matter? an experimental approach to evaluate power market modeling choices (2020). arXiv:2010.16142.

[20] T. Mai, J. Bistline, Y. Sun, W. Cole, C. Marcy, C. Namovicz, D. Young, The role of input assumptions and model structures in projections of variable renewable energy: A multi-model perspective of the u.s. electricity system, Energy Economics 76 (2018) 313–324. doi:https://doi.org/10.1016/j.eneco.2018.10.019.
URL https://www.sciencedirect.com/science/article/pii/S0140988318304213

[21] R. C. Pietzcker, F. Ueckerdt, S. Carrara, H. S. de Boer, J. Després, S. Fujimori, N. Johnson, A. Kitous, Y. Scholz, P. Sullivan, G. Luderer, System integration of wind and solar power in integrated assessment models: A cross-model evaluation of new approaches, Energy Economics 64 (2017) 583–599. doi:https://doi.org/10.1016/j.eneco.2016.11.018.
URL https://www.sciencedirect.com/science/article/pii/S0140988316303395

[22] V. Bosetti, G. Marangoni, E. Borgonovo, L. Diaz Anadon, R. Barron, H. C. McJeon, S. Politis, P. Friley, Sensitivity to energy technology costs: A multi-model comparison analysis, Energy Policy 80 (2015) 244–263. doi:https://doi.org/10.1016/j.enpol.2014.12.012.
URL https://www.sciencedirect.com/science/article/pii/S0301421514006776

[23] J. T. Wilkerson, B. D. Leibowicz, D. D. Turner, J. P. Weyant, Comparison of integrated assessment models: Carbon price impacts on u.s. energy, Energy Policy 76 (2015) 18–31. doi:https://doi.org/10.1016/j.enpol.2014.10.011.
URL https://www.sciencedirect.com/science/article/pii/S0301421514005552

[24] J. Aldy, W. Pizer, M. Tavoni, L. A. Reis, K. Akimoto, G. Blanford, C. Carraro, L. E. Clarke, J. Edmonds, G. C. Iyer, H. C. McJeon, R. Richels, S. Rose, F. Sano, Economic tools to promote transparency and comparability in the paris agreement, Nature Climate Change 6 (2016) 1000–1004. doi:https://doi.org/10.1038/nclimate3106.
URL https://www.nature.com/articles/nclimate3106#citeas

[25] V. Krey, F. Guo, P. Kolp, W. Zhou, R. Schaeffer, A. Awasthy, C. Bertram, H.-S. de Boer, P. Fragkos, S. Fujimori, C. He, G. Iyer, K. Keramidas, A. C. Köberle, K. Oshiro, L. A. Reis, B. Shoai-Tehrani, S. Vishwanathan, P. Capros, L. Drouet, J. E. Edmonds, A. Garg, D. E. Gernaat, K. Jiang, M. Kannavou, A. Kitous, E. Kriegler, G. Luderer, R. Mathur, M. Muratori, F. Sano, D. P. van Vuuren, Looking under the hood: A comparison of techno-economic assumptions across national and global integrated assessment models, Energy 172 (2019) 1254–1267. doi:https://doi.org/10.1016/j.energy.2018.12.131.
URL https://www.sciencedirect.com/science/article/pii/S0360544218325039

[26] S. Giarola, S. Mittal, M. Vielle, S. Perdana, L. Campagnolo, E. Delpiazzo, H. Bui, A. A. Kraavi, A. Kolpakov, I. Sognnaes, G. Peters, A. Hawkes, A. C. Köberle, N. Grant, A. Gambhir, A. Nikas, H. Doukas, J. Moreno, D.-J. van de Ven, Challenges in the harmonisation of global integrated assessment models: A comprehensive methodology to reduce model response heterogeneity, Science of The Total Environment 783 (2021) 146861. doi:https://doi.org/10.1016/j.scitotenv.2021.146861.
URL https://www.sciencedirect.com/science/article/pii/S0048969721019318

[27] ENTSO-E: TYNDP - market modelling data (2016) [cited 21.10.2019].
URL https://www.entsoe.eu/Documents/TYNDP%20documents/TYNDP%202016/rgips/TYNDP2016%20market%20modelling%20data.xlsx

[28] H. C. Gils, Balancing of Intermittent Renewable Power Generation by Demand Response and Thermal Energy Storage, Ph.D. thesis, Universität Stuttgart (2015). doi:10.18419/opus-6888.

[29] Alexander Zerrahn, Wolf-Peter Schill, Long-run power storage requirements for high shares of renewables: review and a new model, Renewable and Sustainable Energy Reviews 79 (2017) 1518–1534. doi:10.1016/j.rser.2016.11.098.

[30] C. Gaete-Morales, M. Kittel, A. Roth, W.-P. Schill, DIETERpy: a Python framework for The Dispatch and Investment Evaluation Tool with Endogenous Renewables, SoftwareX 15 (2021) 100784. doi:10.1016/j.softx.2021.100784.

[31] Ninghong Sun, Modellgestützte Untersuchung des Elektrizitätsmarktes: Kraftwerkseinsatzplanung und -Investitionen, Ph.d. thesis, Institute of Energy Economics and Rational Energy Use, University of Stuttgart (2013). doi:10.18419/opus-2159.

[32] L. Torralba-Díaz, C. Schimeczek, M. Reeg, G. Savvidis, M. Deissenroth-Uhrig, F. Guthoff, B. Fleischer, K. Hufendiek, Identification of the efficiency gap by coupling a fundamental electricity market model and an agent-based simulation model, Energies 13 (15) (2020). doi:10.3390/en13153920.
URL https://www.mdpi.com/1996-1073/13/15/3920

[33] Christian Bußar, Untersuchung optimaler transformationspfade bis 2050 für die erfolgreiche umsetzung einer nachhaltigen reduzierung der kohlendioxidemissionen im bereich der stromerzeugung, Ph.D. thesis, RWTH Aachen University (2019). doi:10.18154/RWTH-2019-09975.

[34] Marius Siemonsmeier, Niklas Bracht, Christian Bußar, Transformation des energiesystems mit steigendem anteil erneuerbarer energien mit netz- und

speicherausbau unter einer gesamteuropäischen perspektive; thema: Entwicklung eines simulationsprogramms und untersuchung von energieversorgungsszenarien : Schlussbericht : Laufzeit: 2014-2017 (2018). doi:10.2314/KXP:1687958572.

[35] Felix Böing, Alexander Murmann, Christoph Pellinger, Stephan Kigle, ISAaR - Integrated simulation model for unit dispatch and expansion with regionalization, Tech. rep., FfE (2019).
URL https://www.ffe.de/en/isaar

[36] Christoph Pellinger, Mehrwert Funktionaler Energiespeicher aus System- und Akteurssicht, Ph.D. thesis, Technische Universität München, https://mediatum.ub.tum.de/1303981 (2016).
URL http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20161012-1303981-1-0

[37] Felix Böing, Cross-sector assessment of CO2 abatement measures and their impact on the transmission grid, Ph.D. thesis, Technische Universität München, https://mediatum.ub.tum.de/1539536 (2020).
URL https://mediatum.ub.tum.de/doc/1539536/1539536.pdf

[38] Peter Meibom, Rüdiger Barth, Bernhard Hasche, Heike Brand, Christoph Weber, Mark O'Malley, Stochastic optimization model to study the operational impacts of high wind penetrations in ireland, IEEE Transactions on Power Systems 26 (3) (2011) 1367–1379. doi:10.1109/TPWRS.2010.2070848.

[39] Tim Drees, Simulation des europäischen Binnenmarktes für Strom und Regelleistung bei hohem Anteil erneuerbarer Energien, Ph.D. thesis, RWTH Aachen (2015).
URL https://publications.rwth-aachen.de/record/658687

[40] Simon Hilpert, Cord Kaldemeyer, Uwe Krien, S. Günther, Clemens Wingenbach, Guido Pleßmann, The open energy modelling framework (oemof) - a new approach to facilitate open science in energy system modelling, Energy Strategy Reviews 22 (2018) 16–25. doi:10.1016/j.esr.2018.07.001.
URL https://www.sciencedirect.com/science/article/pii/S2211467X18300609

[41] Uwe Krien, Patrik Schönfeldt, Jann Launer, Simon Hilpert, Cord Kaldemeyer, Guido Pleßmann, oemof.solph—a model generator for linear and mixed-integer linear optimisation of energy systems, Software Impacts 6 (2020) 100028. doi:10.1016/j.simpa.2020.100028.
URL https://www.sciencedirect.com/science/article/pii/S2665963820300191

[42] oemof.solph documentation.
URL https://oemof-solph.readthedocs.io/en/latest/

[43] Hans Christian Gils, Yvonne Scholz, Thomas Pregger, Diego Luca de Tena, Dominik Heide, Integrated modelling of variable renewable energy-based power supply in Europe, Energy 123 (2017) 173–188. doi:10.1016/j.energy.2017.01.115.